MARS EXPLORATION PROGRAM

DATA MANAGEMENT PLAN

Raymond E. Arvidson
Susan Slavney
Stephanie Nelson
Dept. of Earth and Planetary Sciences
Washington University
St. Louis, MO 63130

Rev. 3.0
March 20, 2002

Mars Exploration Program
Data Management Plan

Approved by:

Elaine Dobinson **Date**
Project Manager
Planetary Data System

James Garvin **Date**
Program Scientist
Mars Exploration Program, NASA Headquarters

Orlando Figueroa **Date**
Program Director
Mars Exploration Program, NASA Headquarters

Joseph H. King **Date**
National Space Science Data Center

Daniel J. McCleese **Date**
Program Scientist
Mars Exploration Program, Jet Propulsion Laboratory

Firouz Naderi **Date**
Program Manager
Mars Exploration Program, Jet Propulsion Laboratory

Table of Contents

Changes iv

Acronyms v

1 Purpose and Scope of Document 1

2 Overview of the Mars Exploration Program 1

3 Overview of the Planetary Data System 1

4 Roles and Responsibilities 2

4.1 Mars Exploration Program 2
4.2 Planetary Data System 3
4.3 National Space Science Data Center 3
4.4 Jet Propulsion Laboratory Media Relations Office 3
4.5 Principal Investigator Home Institution Media Relations Offices 3

5 Data Products 3

5.1 Packetized Data Records 4
5.2 Experiment Data Records 4
5.3 Standard Data Products 4
5.4 Ancillary Files 4
5.4.1 SPICE Files 4
5.4.2 Engineering and Other Data 5
5.5 Public Information Products 5
5.6 Data Set Documentation 5
5.7 Software 6

6 Policies for Release of Data and Public Information 6

6.1 Data Rights and Release Policy 6
6.2 Public Information Release Policy 7
6.3 Data Release Acknowledgement Policy 7

7 Documentation Concerning Returned Samples 7

8 Archive Planning, Generation, Validation, and Delivery 8

8.1 Archive Media 8
8.2 Planning and Documentation 8
8.3 Generation 9
8.4 Validation and Delivery 9

9 Applicable Documents 10

10 Definitions of Terms 11

11 Tables 12

12 Figures 20

Tables

Table 1 – Instruments on Current MEP Projects and Expected Instruments on Upcoming Projects. 12
Table 2 – Example Outline for Project Archive Generation, Validation and Distribution Plan 13
Table 3 – NASA Definition of Processing Levels for Science Data Sets.. 14
Table 4 – Annotated Example Outline for Data Product SIS for Instrument-Related Data 15
Table 5 – Example Outline for Archive Volume SIS for Instrument-Related Data 18
Table 6 – Example of PDS Archive Volume Structure and Contents 19

Figures

Figure 1 – Timeline for Exploration Program Projects. 20
Figure 2 – Planetary Data System Organizational Structure. 21
Figure 3 – Flow of Archive Volume Generation, Validation and Distribution Functions for Mars Exploration Program Projects.. 22

Changes

Date	Section	Change
5/20/99	6.1, Data Rights and Release Policy	Added paragraph VII
5/20/99	6.3, Data Release Acknowledgement Policy	New section
5/28/99	6.2, Public Information Release Policy	Added paragraph VI
6/21/99	6.1, Data Rights and Release Policy	Paragraph VII: added “immediately”
6/21/99	6.2, Public Information Release Policy	Paragraph VI: added “immediately”
6/21/99	2, Overview of the Mars Surveyor Program; Signature page; Acronyms	Removed “Mars Exploration Directorate”; added “Space Science Exploration Division”.
8/26/99	Signature page	Replaced Thomas E. Thorpe with Richard Cook. Added Chris Jones. Changed title for Daniel J. McCleese.
9/13/99	Signature page	Added Joseph Boyce.
4/11/00	6.1, Data Rights and Release Policy; 7, Documentation Concerning Returned Samples	Incorporated several edits by J. Boyce
4/11/00	Signature page	Changed PDS Project Manager from Rosana Borgen to Elaine Dobinson
11/2/00	6.1, Data Rights and Release Policy; 7, Documentation Concerning Returned Samples	Removed 4/11/00 edits by J. Boyce
11/2/00	All	Changed "Mars Surveyor Program" to "Mars Exploration Program"; removed references to MSOP; updated references to current and future missions
11/2/00	Signature page	Replaced Mars Surveyor Program personnel with Mars Exploration Program personnel
1/22/02	Signature page	Updated names
1/22/02	Table 1	Added MRO Specifications
1/23/02	Acronyms	Deleted MECA, MIP; Added IDS
1/28/02	Figure 1	Corrected mission phase lengths
1/28/02	Figure 3	Revised data flow
1/28/02	Table 4, 5	Updated outlines
1/28/02	All	Responded to review comments
1/29/02	Section 5.6, Data Set Documentation	Added PDS Catalog file responsibilities
3/20/02	Signature page	Added Garvin and McCleese

Acronyms

APXS	Alpha-Particle X-ray Spectrometer
ASI/MET	Atmospheric Structure Instrument/Meteorology Package
CD-ROM	Compact disc read only memory
CODMAC	Committee on Data Management and Computation
DAWG	Data and Archive Working Group
DSN	Deep Space Network
EDR	Experiment Data Record
GRS	Gamma Ray Spectrometer
IMP	Imager for Mars Pathfinder
IR	Infrared
JPL	Jet Propulsion Laboratory
MAG/ER	Magnetometer / Electron Reflectometer
MARCI	Mars Color Imager
MARIE	Mars Radiation Environment Experiment
MEP	Mars Exploration Program
MFEX	Microrover Flight Experiment (AKA: Rover, Sojourner)
MGS	Mars Global Surveyor
MOC	Mars Orbiter Camera
MOLA	Mars Orbiter Laser Altimeter
MPF	Mars Pathfinder
NAIF	Navigation and Ancillary Information Facility
NASA	National Aeronautics and Space Administration
NSSDC	National Space Science Data Center
PDS	Planetary Data System
PMIRR	Pressure Modulator Infrared Radiometer
PSG	Project Science Group
RDR	Reduced Data Records
SDVT	Science Data Validation Team
SIS	Software Interface Specification
SPICE	Spacecraft, Planet, Instrument, C-matrix, Events
TES	Thermal Emission Spectrometer

1 Purpose and Scope of Document

The Mars Exploration Program (MEP) Data Management Plan presents a high-level plan for timely generation, validation, and delivery of data products from MEP projects to the Planetary Data System (PDS) and the National Space Science Data Center (NSSDC) in complete, well-documented, permanent archives. The Plan also specifies policies and procedures for distributing the data and information to the general public in a timely fashion.

This document is for use by personnel associated with MEP projects to understand data release policies and how data products are to be archived. In fact, each Project will write an Archive Plan that is responsive to this document. Principal Investigators associated with MEP projects will use this document to understand what is expected of them in terms of delivery of a complete data archive, and how the PDS can help them prepare their data products. PDS and NSSDC personnel will use this document to understand the types of data products and archives they can expect to receive from MEP projects, in order to allocate resources accordingly.

2 Overview of the Mars Exploration Program

MEP is a multiyear effort with the intent of implementing a coordinated exploration of Mars. The focus is on understanding current and past climates, searching for evidence of past and extant life, and determining the extent to which human expeditions to the planet are feasible.

Implementation consists of launches during opportunities approximately 2-years apart (Figure 1). Mars Global Surveyor was the first project, followed by the 2001 Odyssey Orbiter. Future opportunities will include a pair of Mars Exploration Rovers to be launched in 2003 and the Mars Reconnaissance Orbiter in 2005. Scouts, orbiters, rovers, and sample return missions form the template for opportunities beyond 2005.

Science payloads for approved MEP Projects are listed in Table 1. The long-term scientific payoff of MEP will come about if the observations and resultant data feed from one project to the next and if documented archives are used by the scientific community to increase knowledge about the planet.

3 Overview of the Planetary Data System

The NASA PDS is an active archive that provides high quality, usable planetary science data products to the science community (Arvidson and Dueck, 1994, McMahon, 1996). This system evolved in response to science community requests for improved availability of planetary data from NASA projects through increased scientific involvement and oversight. PDS provides access to data archives for scientists, educators, and the public.

The first objective of the PDS is to publish and disseminate documented data sets of use in scientific analyses. Whereas the media of the published data vary, all PDS-produced products are reviewed by scientists and data engineers to ensure that the data and the related materials are appropriate and usable. PDS data sets are typically published as archives, collections of reviewed data and documentation, ancillary information (such as calibration data), software, and any other tools needed to understand and use the data. PDS provides access to data by a system of online and compact disc archives.

The second objective of the PDS is to work with projects to help them design, generate, and validate their data products as archives. The advantage of having projects deliver well-documented products is that use by the scientific and public communities is maximized.

The third PDS objective is to develop and maintain archive data standards to ensure future usability. PDS has developed a set of standards for describing and storing data so that future scientists unfamiliar with the original experiment can analyze the data using a variety of computer platforms with no additional support beyond that provided with the product. These standards address structure of the data, description of contents, media design, and a standard set of terms (PDS Standards Reference, 1995; Planetary Science Data Dictionary, 1996).

The fourth PDS objective is to provide expert scientific help to the user community. Most of the archive products may be accessed or ordered automatically by users, but PDS also provides teams of scientists to work with users to select and understand data. In some cases, special processing can be done to generate user-specific products. In addition, PDS-funded scientists provide a sounding board to the science community to for comments on the need for new and/or improved software, data sets, and archives.

To accomplish the four objectives, the PDS is structured as a distributed system composed of eight teams, called nodes (Figure 2). The science nodes include the Atmospheres, Geosciences, Planetary Plasma Interactions, Rings, and Small Bodies Nodes. Each has an advisory group of discipline scientists to provide guidance and priorities for PDS. The support nodes are the Imaging Node, with expertise in sophisticated image processing and cartography, the Navigation and Ancillary Information Facility (NAIF) Node, specializing in observation geometry and sequence of events information, and the Central Node, leading the project, at the Jet Propulsion Laboratory.

4 Roles and Responsibilities

This section summarizes the roles and responsibilities for personnel and organizations involved in the generation, validation, transfer, and distribution of MEP archives.

4.1 Mars Exploration Program

Each MEP project is responsible for the production and delivery to PDS of documented archives that use PDS standards. Instrument Principal Investigators are responsible, under contractual obligation, for working with PDS personnel on planning, generating, and validating archives from their instruments. The Project Science Group (PSG) will coordinate plans for archiving through such entities as a Data and Archive Working Group (DAWG) and the Science Data Validation Team (SDVT). Archives must include at least three copies on hard digital media produced under Project auspices and destined for long-term storage at PDS and NSSDC sites, in addition to any on-line transfers.

Each MEP project will generate an archive plan document that describes the specific generation, validation, and transfer to PDS of the project's data products. The plan will specify the contents of archives, the personnel or organization responsible for generating each archive, estimates of the amount of data to be delivered, and a schedule for delivery. The Project archive plans will be based on the policies and procedures specified in the Mars Exploration Program Data Management Plan (this document). An outline of a project archive plan is provided in Table 2.

Each MEP project is responsible for distributing its data products internally, i.e. among scientists associated with the project. If this distribution takes the form of a mass production of validated archive volumes (e.g. CD-ROMs), the PDS may ask to fund a number of additional copies for its own distribution, thereby saving production setup costs.

The Project Office of each MEP project is responsible for ensuring that the policies and implementation plans concerning release of information and interactions with the public and press are followed. The Project also is responsible for primary interactions with NASA Headquarters, including ensuring that Headquarters personnel are aware of information releases before they occur, and focusing feedback from Headquarters back to the appropriate parties. The Project has the responsibility for primary interactions with the JPL Media Relations Office, serving as coordinating entity for media relations focused at JPL.

The Project Scientist has primary responsibility for ensuring that scientific findings are made available to the media. In addition, the Project Scientist or his/her designates, working with the JPL Media Relations Office, will:

a. Help instrument teams prepare plans for generation and release of public-oriented data and results.

b. Ensure that PSG members prepare release materials with enough lead time so that NASA Headquarters receives the materials about a week before the general release to the media.

c. Coordinate the preparation of materials (press releases, interviews, press conference scripts, video file productions, etc.) with the JPL media relations representative.

d. Identify themes or subjects for each media interaction opportunity months in advance, and to schedule the preparation work.

e. Decide what the best media interaction vehicle should be for the information to be presented.

f. Help enforce the graphical standards of the Media Relations Office for best television viewing.

4.2 Planetary Data System

The PDS is the primary organization within NASA responsible for archiving and distributing planetary data. The PDS is thus the designated point of contact for MEP projects on archive-related issues. The PDS is also the interface between MEP projects and the NSSDC.

The PDS Geosciences Node Director has been selected to take the lead in working with each MEP project to ensure that overall plans are in place and that individual projects are meeting archiving commitments. Relevant PDS personnel will participate in archive planning efforts through DAWGs and the SDVT for each project to ensure that: (a) archives are planned, generated, and reviewed using PDS standards, (b) archives are constructed in ways that facilitate cross-mission and cross-instrument data analyses, (c) PDS provides user services that allow search and access to data in ways that transcend particular instruments or data sets.

The PDS is responsible for distributing MEP project archives to the broad science community after the data have been released by the project. Distribution may be done electronically via the Internet or by replication of archive volumes for delivery to users. The PDS maintains copies of project archives, and the relevant PDS discipline nodes are available to provide information and expert assistance to users of the data.

4.3 National Space Science Data Center

The NSSDC's primary role is to maintain a "deep archive" of data for long-term preservation, for filling any large orders requested by the science community, and for filling orders requested of the NSSDC by the general public. The PDS will deliver at least one hard medium digital copy of MEP project archive volumes to NSSDC.

4.4 Jet Propulsion Laboratory Media Relations Office

The JPL Media Relations Office is chartered to be the prime point of contact with the public and the media for activities focused at JPL. The overall objective of the Media Relations Office is to maximize the positive impact of information releases from the project. This office helps organize and implement press releases and press conferences. They facilitate release of data and information meant for consumption by the press and public, including video feeds of interviews.

4.5 Principal Investigator Home Institution Media Relations Offices

Each Principal Investigator may release findings from his or her instrument to the press, through the home institution media relations office, if those releases are coordinated with the MEP Project Scientist and the JPL Media Relations Office.

5 Data Products

The science-related data products to be produced during MEP projects are described in this section, including standard products, engineering and other ancillary products, and documentation that accompany data sets as they are delivered to the PDS. For reference, Table 3 defines numerical processing levels, using the common assignments developed by the NASA community.

5.1 Packetized Data Records

These data records consist of time-ordered packetized telemetry received from the spacecraft, with duplications removed, together with ancillary information needed to understand what is contained in a given packet.

5.2 Experiment Data Records

Experiment Data Records (EDRs) are NASA Level 0 products (see Table 3) and consist of time-ordered sequences of raw science data obtained by a given instrument, together with engineering information that allows instrument teams to check operation of their instruments. The specific contents of EDR data will be instrument-dependent.

5.3 Standard Data Products

Standard Data Products include both EDRs and Reduced Data Records (RDRs). RDRs are typically generated from EDRs and SPICE data at investigator facilities. RDRs are defined during the proposal and selection process and are contractually promised by the Principal Investigators as part of the investigation. Standard Data Products are produced in a systematic way during the course of the project, primarily by Instrument Teams, although some may be generated by Interdisciplinary Scientists addressing specific scientific problems. Standard Data Products are sometimes called operational products in other projects, since they are produced routinely with fairly well-defined procedures. During MEP projects it is expected that increased knowledge of instrument calibrations and operating nuances, together with increased knowledge of Mars, and better understanding of how to reduce raw data, will result in updates to some Standard Data Products.

Standard Data Products shall be delivered to the PDS in PDS-compatible format, including PDS labels and documentation. The documentation shall be thorough enough to understand quantitatively the processing history.

5.4 Ancillary Files

5.4.1 SPICE Files

SPICE is an acronym used to describe a suite of elemental ancillary data sets, often called kernels:

- S Spacecraft Ephemeris, with files containing spacecraft location as a function of time
- P Planetary/Satellite ephemerides and associated target body physical and cartographic constants
- I Instrument information, including mounting alignment and field-of-view size/shape/orientation
- C Orientation of a spacecraft's primary coordinate systems and possibly angular rates as well
- E Event information, including nominal sequences, real-time commanding, unscheduled events, and experimenter's notebook comments

Most SPICE kernels will be generated by flight operations teams with some instrument information and sequence data obtained from instrument teams. Where appropriate, there will be both predict and reconstructed SPICE files. Reconstructed SPICE kernels will typically be generated within a couple of weeks after data acquisition and made available to instrument teams. SPICE kernels will be used together with a toolkit of software modules (to be supplied by JPL's Navigation and Ancillary Information Facility) at investigator institutions to generate the derived ancillary data needed to help plan observations and to process EDRs (see: NAIF, 1988). SPICE kernels and NAIF toolkit software are applicable to landers and rovers as well as to orbiting spacecraft.

5.4.2 Engineering and Other Data

Engineering data, while not strictly science data, can be used to aid the interpretation of science data and as such will be archived. Other data that may be useful include sequence of events files, log files, DSN monitor files, etc. The specific content of engineering and other ancillary data products will be negotiated between each Project and the PDS and incorporated into the Project Archive Plan. These products will be archived by NAIF or another entity as designated in the future. In some cases, engineering data may be bundled with science archives, e.g. rover wheel currents and rover yaw, pitch, and roll data would be needed to infer soil properties during traverses and trenching operations. Thus, these engineering data might be included in archives focused on imaging and *in-situ* data acquired during rover operations.

5.5 Public Information Products

Public Information Products consist of selected data products that have been annotated for the media with explanations and delivered to the relevant NASA or home institution offices for use in public relations. It is expected that these products will be produced mainly at investigator institutions, although hard copies for press distribution may be produced at JPL.

5.6 Data Set Documentation

Documentation of data acquisition and processing histories is crucial to successful long-term use of project data. The SPICE E Kernel will contain information delineating why data sets were acquired, including whether they were part of a coordinated sequence. Information on spacecraft events that affected the sequences and data products will also be included. The actual (as opposed to predict) SPICE files will be transferred to a long-term archive along with raw and reduced data.

Information needed to fully understand EDRs, SPICE files, and reduced data will also be generated for long-term use of MEP project data. Documentation to be provided by each instrument team will typically include:

- Calibration Requirements Document
- Calibration Report
- In-flight Calibration Report
- Science Reports
- User's Guides that describe data products from particular instruments and how they were generated
- Software User's Guides that describe processing software delivered as part of the archive
- Instrument-specific sequence of events, in E-kernel format

Documentation to be provided by each Project will include a Mission Operations Report, stipulating how the mission plan was actually carried out.

Data sets archived with PDS are described by entries in the PDS Catalog. Data set providers are required to supply descriptive information for the PDS Catalog. Specifically, the Project is responsible for descriptions of the mission and the instrument hosts (spacecraft), and instrument teams are responsible for descriptions of their instruments and data sets.

5.7 Software

It is expected that standard data products delivered by Principal Investigators should be in a format that is accessible by other users. If the products cannot be accessed using commonly available software tools, then the products should be accompanied by software or detailed algorithms that allow a user to access the data.

6 Policies for Release of Data and Public Information

6.1 Data Rights and Release Policy

Because of the expected widespread scientific and public interest in new results from Mars and the strong MEP commitment to releasing data on a timely basis, it is important to establish a clear release policy for archives and other data sets (e.g. quick-look products).

A key element of the policy is the need for a reasonable interval of time to generate and validate standard data products and archive volumes before release to the general community. Based on numerous past experiences (including the Magellan, Clementine, and MGS projects), up to a six-month period is necessary to produce useful RDRs for data from most instruments. EDRs and RDRs should be produced, validated, and released as soon as possible, with the six-month period considered to be a maximum (although certain data sets such as global maps may require more time to generate).

Specific release policies are:

I. The generation/validation time period for standard data products is defined to be the period from receipt of science packets containing raw data at instrument processing facilities until release of archive volumes to the PDS. The archive volumes will include, as appropriate, standard data products, SPICE files, relevant software, and documentation describing the generation of the products. During the generation/validation period, Instrument Teams are expected to use raw data to generate standard products and archives, and to validate the products and volumes. Efforts involving all of the project investigators are expected to be underway during this period, including product validation. Generation and validation of products and archives may typically require up to six months from the time of receipt of raw data at instrument processing facilities in order to produce and validate useful archive volumes. More rapid release is highly desired.

II. To ensure rapid dissemination of new and significant information, each Principal Investigator will also release a significant subset of data earlier as a form of public outreach and education. These releases will typically be available within a week of data receipt. Postings on the Internet (e.g., World Wide Web) will be used as a cost-effective way for widespread dissemination of these special products. The posted data may include images, derived spectra, topographic information, and other forms of data that illustrate new and significant results. Postings will include documentation. These data releases will conform to the Public Information Release Policy (section 6.2 below).

III. During the generation/validation period, use and analysis of raw and derived products from a particular instrument, use of the results of unpublished papers derived from such analysis, or use of data posted on the Internet for public outreach and education should be done only with the agreement of the relevant Principal Investigator.

IV. After the six-month or less generation/validation period, the relevant archive volumes will be transferred to the PDS, which will make them available to the general science community.

V. Individual Instrument Teams are strongly encouraged to make data available as soon as possible, and they may decide to make a number of versions available to the PDS on a more timely basis than the typical six-month period. They may not release data from other instruments.

VI. Items I. to V. above pertain to nominal mission operations. In the event of anomalies, such as the prolonged aerobraking phase for MGS, data should be archived using stated policies, unless observational, personnel, and/or financial constraints force the need for a longer period between data receipt and transfer of archives to the PDS.

VII. MEP reserves the right to require a Principal Investigator to acquire and/or release immediately a limited set of data for purposes of mission planning and ensuring mission safety, for both current and future missions.

6.2 Public Information Release Policy

Public information release includes press conferences and written material concerning both mission operations and scientific analyses. Specific policy statements for Public Information Release for Mars Exploration Program projects are:

I. Information concerning spacecraft and instrument anomalies may be released only by the Project Office, in coordination with NASA Headquarters and Principal Investigators.

II. Information concerning significant scientific results may be released during press conferences, press releases organized by the Project Office and the Jet Propulsion Laboratory's Media Relations Office, and on the World Wide Web, in coordination with NASA Headquarters and relevant members of the project's science teams.

III. Information concerning scientific results from a given instrument may also be released by the Principal Investigator for the instrument. For example, such releases may be organized by the home institution of the Principal Investigator. These releases can be in any medium, but significant results must have prior coordination through the JPL Project Office and NASA Headquarters.

IV. Information concerning scientific results from instrument studies may be released by an Interdisciplinary Scientist, Participating Scientist, or Co-Investigator before it is deposited in the Planetary Data System, only with the concurrence of the Principal Investigator whose instrument data are used.

V. In cases II through IV the Project Office must receive, in advance of the release, a copy of the release material (e.g., images, spectra, captions, summary of results), a schedule for the release, and a statement of the mechanisms for release. The intent is not to require concurrence for the release, but only to make sure that the Project Office, the Public Information Office, and NASA Headquarters are informed of the releases before they happen.

VI. MEP reserves the right to require a Principal Investigator to acquire and/or release immediately a limited set of data for purposes of public or educational outreach.

6.3 Data Release Acknowledgement Policy

In any release of data or information as described in sections 6.1 and 6.2 above, appropriate credit will be given to all entities involved, including the Principal Investigator, the Principal Investigator's home institution, JPL, and NASA. The credit will appear on public web sites that contain data releases and on documentation that accompanies the data products.

7 Documentation Concerning Returned Samples

Samples of rocks and soils may be returned from Mars as part of MEP. It is expected that the samples will be archived in an appropriate curatorial facility or facilities, where initial analyses will be conducted. This facility or facilities will also serve as a source of sample material for the science community.

It is crucial that observations that provide context (e.g., geological, geochemical, environmental) for returned samples also be archived. Such data must be archived in ways that can be easily linked to sample information developed by the curatorial facilities. To facilitate such links, it is expected that the PDS, MEP, and curatorial facility personnel will work together to establish data bases and procedures to connect mission and sample archives.

8 Archive Planning, Generation, Validation, and Delivery

Standard products form the core of the archives to be produced by MEP projects and released to the PDS for distribution to the science community. These products, SPICE files, software, and engineering and other ancillary information will be validated for transfer to the PDS.

The following section discusses the processes and schedules for generation and validation of standard products and archives by MEP projects. Figure 3 shows the flow of data through the stages of archives generation, validation, transfer to the Planetary Data System, and distribution to the science community. Also shown is posting of timely results on the Internet for education and outreach.

8.1 Archive Media

An *archive* consists of one or more data sets along with all the documentation and ancillary information needed to understand and use the data. Thus defined, an archive is a logical construct independent of the medium on which it is stored.

An *archive volume* is a unit of media on which an archive is stored; for example, one CD-ROM. When an archive spans multiple volumes, they are called an *archive volume set*. Usually the documentation and some ancillary files are repeated on each volume of the set, so that a single volume can be used alone.

Whole or partial archives may be placed online for convenient distribution via the Internet. However, to ensure long-term viability, archives must be also stored offline on physical archive volumes. The choice of storage medium will depend on available technology, but it should be a stable, industry-standard medium intended for long-term storage. At present, the standard media for long term storage are CD-ROMs and DVD-ROMs. Note that it is a requirement for each project to transfer to the PDS three copies of archives on hard digital media.

8.2 Planning and Documentation

Preparation for generating project data archives involves designing the data products, the online systems, and the physical archive volumes on which they will be stored, and developing the systematic generation, validation, and distribution procedures for these archive volumes. These plans will be documented in an Archive Volume Generation, Validation and Distribution Plan created by each project. These documents will cover all the standard data products generated by the project (e.g., Arvidson et al., 1997; Arvidson, 2001; Arvidson et al., 2002a; Arvidson et al., 2002b). For reference, Table 2 shows a suggested outline for an Archive Plan.

The format and content of data products are described in Data Product Software Interface Specification (SIS) documents, created by the instrument Principal Investigators or ancillary data product's cognizant engineer. Typically a separate SIS is created for each type of data product, including raw, derived and ancillary products. Table 4 is an example of a Data Product SIS outline, including annotations concerning the specific content recommended for each section. (Note that the example outline given is that of a completely self-contained Data Product SIS. If the pertinent information is available from other sources, e.g. data product labels and catalog descriptions, then sections of the Data Product SIS may simply refer the reader to those sources.)

The format and content of the archive volumes are described in an Archive Volume SIS for each data set, which is also generated by the Principal Investigator for the instrument or a cognizant engineer, working with relevant PDS personnel as needed. Table 5 is an example outline for an Archive Volume SIS. Table 6 shows a typical directory structure and contents for a PDS archive.

8.3 Generation

Generation of science data packets, EDRs, RDRs, and other data sets for a given MEP project is the responsibility of that project, as discussed in Section 4. Generation of archive volumes of raw (level 0) radio science data is the responsibility of the Radio Science Team associated with a given project or the project, as appropriate.

The DAWG or its equivalent for each project will provide oversight on implementation of the Archive Plan for that Project and ensure timely generation, validation, and delivery of archives to the PDS.

The MEP SDVT will provide overall oversight on archiving and operational processes for all MEP Projects, focusing on inter-project issues such as distributed operations infrastructures and related matters.

The Navigational and Ancillary Information Facility (NAIF) will generate an archive volume set of all SPICE data.

8.4 Validation and Delivery

Validation of engineering and science packets and SPICE files will be an intrinsic part of standard product generation. Further science validation of standard products will be done in part during analysis of the data.

A key additional requirement is the validation of archives for integrity of scientific content, file structures (e.g., do the files conform to Software Interface Specification documents?), directory structures, compliance with PDS standards, and integrity of the physical media used to transfer the data products. Validation will be coordinated through the DAWG or its equivalent and will follow the relevant Archive Plan. Interested members of the scientific community may also be asked to comment on plans for archive generation, validation, and delivery.

During the validation period, problems due to obvious errors in processing, missing files, and inadequate documentation will be referred back to the data suppliers for correction. Minor errors may simply be documented on the volumes. A fundamental aspect of the release schedule is that science packets, SPICE files, and algorithms/software generating derived data products are released at the same time as reduced data records generated from the relevant science packet data.

The DAWG will periodically report results of validation to the Project Science Group for the specific project. If the volumes are approved for release by the Project, then the Principal Investigator or relevant Project personnel would transfer the archives to the PDS, based on the release schedule specified by the Project's Archive Generation, Validation, and Distribution Plan. This may include online transfers and must include delivery of three copies of archives on hard digital media.

The PDS is responsible for delivering a copy of each archive volume set to the NSSDC.

9 Applicable Documents

Arvidson, R. E., and S. L. Dueck, 1994, The Planetary Data System, *Remote Sensing Rev.*, **9**, 255-269.

Arvidson, R., E. Guinness, S. Slavney, and R. J. Springer, 1998, *Mars Global Surveyor Project Archive Generation, Validation, and Transfer Plan*, MGS Project Document 542-312.

Arvidson, R., E., 2001, *2001 Mars Odyssey Orbiter Archive Generation, Validation, and Transfer Plan*, Odyssey Project Document 722-302, Rev. 0.

Arvidson, R., S. Slavney, and S. Nelson, 2002a, *Mars Exploration Rover (MER) Project Archive Generation, Validation, and Transfer Plan*, MER Project Document 420-1-200, Draft.

Arvidson, R., and S. Nelson, 2002b, *Mars Reconnaissance Orbiter (MRO) Project Archive Generation, Validation, and Transfer Plan*, MRO Project Document xxx-xxx, Draft.

Data Management and Computation, Volume 1, Issues and Recommendations, 1982, National Academy Press, 167 p.

Issues and Recommendations Associated with Distributed Computation and Data Management Systems for the Space Sciences, 1986, National Academy Press, 111 p.

McMahon, S. K., 1996, Overview of the Planetary Data System, *Planet. Space Sci.*, **44**, 3-12.

Planetary Science Data Dictionary Document, July 15, 1996, Planetary Data System, JPL D-7116, Rev. D.

Planetary Data System Data Preparation Workbook, February 17, 1995, Version 3.1, JPL D-7669, Part-1.

Planetary Data System Standards Reference, July 24, 1995, Version 3.2. JPL D-7669, Part-2.

10 Definitions of Terms

Archive – An archive consists of one or more data sets along with all the documentation and ancillary information needed to understand and use the data. An archive is a logical construct independent of the medium on which it is stored.

Archive Volume, Archive Volume Set – A volume is a unit of media on which data products are stored; for example, one CD-ROM. An *archive volume* is a volume containing all or part of an archive; that is, data products plus documentation and ancillary files. When an archive spans multiple volumes, they are called an *archive volume set*. Usually the documentation and some ancillary files are repeated on each volume of the set, so that a single volume can be used alone.

Data Product – A labeled grouping of data resulting from a scientific observation, usually stored in one file. A product label identifies, describes, and defines the structure of the data. An example of a data product is a planetary image, a spectrum table, or a time series table.

Data Set – An accumulation of data products. A data set together with supporting documentation and ancillary files is an archive.

Experiment Data Records – NASA Level 0 data for a given instrument; raw data.

Reduced Data Records – Science data that have been processed from raw data to NASA Level 1 or higher. See Table 3 for definitions of processing levels.

Standard Data Product – A data product that has been defined during the proposal and selection process and that is contractually promised by the Principal Investigator as part of the investigation. Standard data products are generated in a predefined way, using well-understood procedures, and processed in "pipeline" fashion.

11 Tables

Table 1 – Instruments on Current MEP Projects and Expected Instruments on Upcoming Projects.

Note: Mars Pathfinder is a Discovery-class project. It is included because of relevance to MEP projects and objectives.

Project	**Instruments**
Mars Pathfinder	ASI/MET (Atmospheric Structure Instrument / Meteorology Package) IMP (Imager for Mars Pathfinder) MFEX (Microrover Flight Experiment (Sojourner Rover)) with APXS (Alpha Proton X-Ray Spectrometer)
Mars Global Surveyor	MAG/ER (Magnetometer/ Electron Reflectometer) MOC (Mars Orbiter Camera) MOLA (Mars Orbiter Laser Altimeter) TES (Thermal Emission Spectrometer) RS (Radio Science)
2001 Mars Odyssey Orbiter	THEMIS - Thermal-IR Multispectral Imaging System MARIE - Mars Radiation Environment Experiment GRS - Gamma Ray Spectrometer
2003 Mars Exploration Rovers	Twin rovers with Athena Payload: Pancam Navcam Mini-TES APXS Mössbauer Spectrometer Microscopic Imager Rock Abrasion Tool
2005 Mars Reconnaissance Orbiter	MARCI (Mars Color Imager)
	MCS (Mars Climate Sounder [formerly PMIRR])
	CTX (Mars Context Imager)
	CRISM (Compact Reconnaissance Imaging Spectrometer for Mars)
	HiRISE (High Resolution Imaging Science Experiment)
	SHARAD (Subsurface Sounding Shallow Radar)
	Gravity & Accelerometer Facility Science Investigations

Table 2 – Example Outline for Project Archive Generation, Validation and Distribution Plan

1. Introduction

   1.1 Purpose
   1.2 Scope
   1.3 Contents
   1.4 Applicable Documents and Constraints

2. Overview of Mission

3. Roles and Responsibilities

   3.1 Project
   3.2 Planetary Data System
   3.3 National Space Science Data Center

4. Archive Generation, Validation, Transfer, and Distribution

   4.1 Generation
   4.2 Validation
   4.3 Transfer
   4.4 Distribution

5. Archive Generation, Validation and Release Schedules

Table 1. Payload
Table 2. Definitions of Processing Levels for Science Data Sets
Table 3. Components of Archives
Table 4. Standard Data Products
Table 5. Archive Component Suppliers
Table 6. Archive Generation, Validation, and Release Schedule

Table 3 – NASA Definition of Processing Levels for Science Data Sets

NASA	**CODMAC**	**Description**
Packet data	Raw - Level 1	Telemetry data stream as received at the ground station, with science and engineering data embedded.
Level-0	Edited - Level 2	Instrument science data (e.g., raw voltages, counts) at full resolution, time ordered, with duplicates and transmission errors removed.
Level 1-A	Calibrated - Level 3	Level 0 data that have been located in space and may have been transformed (e.g., calibrated, rearranged) in a reversible manner and packaged with needed ancillary and auxiliary data (e.g., radiances with the calibration equations applied).
Level 1-B	Resampled - Level 4	Irreversibly transformed (e.g., resampled, remapped, calibrated) values of the instrument measurements (e.g., radiances, magnetic field strength).
Level 2	Derived - Level 5	Geophysical parameters, generally derived from Level 1 data, and located in space and time commensurate with instrument location, pointing, and sampling.
Level 3	Derived - Level 5	Geophysical parameters mapped onto uniform space-time grids.

Table 4 – Annotated Example Outline for Data Product SIS for Instrument-Related Data

1. Introduction

1.1. Purpose and Scope of Document

The purpose of this document is to provide users of the data product with a detailed description of the product and a description of how it was generated, including data sources and destinations. The document is intended to provide enough information to enable users to read and understand the data product. The users for whom this document is intended are the scientists who will analyze the data, including those associated with the project and those in the general planetary science community. This section should include a one-sentence description of the data product.

1.2 Contents

This data product SIS describes how the instrument acquires its data, and how the data are processed, formatted, labeled, and uniquely identified. The document discusses standards used in generating the product and software that may be used to access the product. The data product structure and organization is described in sufficient detail to enable a user to read the product. Finally, an example of a product label is provided.

1.3. Applicable Documents

PDS Standards and Data Dictionary documents

Science Data Management Plan for the project

Archive Generation, Validation and Distribution Plan for the project

Archive Volume SIS for the data product

Paper(s) describing the instrument

1.4. Relationships with Other Interfaces

State what products, software, and documents would be affected by a change in this data product. Refer to documents that describe the other interfaces.

2. Data Product Characteristics and Environment

2.1 Instrument Overview

Describe how the instrument acquires data; include a general description of operating modes as they relate to the data product.

2.2 Data Product Overview

Describe the type of data product (image, spectrum, etc.); what it measures (radiance, voltage, time, etc.); how observations are organized into data products; how a data product is stored (binary or ASCII, how grouped into files). Details should be specified in the following sections.

2.3 Data Processing

This section should provide general information about the data product content, format, size, and production rate. Details about data format should be specified later in section 6.

2.3.1 Data Processing Level

Describe the product in terms of its NASA and/or CODMAC processing levels.

2.3.2 Data Product Generation

Describe how and by whom data products are generated. Describe any calibrations, corrections, or compressions that have been applied to the product. Specify software, algorithms, calibration files or procedures. If a product has been compressed or otherwise processed in a reversible manner, indicate the software, algorithms, and/or ancillary data needed to reverse the processing, and describe how they should be applied. State whether multiple versions of the products will be generated.

2.3.3 Data Flow

Describe the sources, destinations, and transfer procedures for data products. State the size of an individual data product and the total size of all the data products generated over the course of each mission phase. State the time span covered by a product, if applicable, and the rate at which products are generated and delivered. If more than one version of a product may be generated, state the number of expected versions and the rate at which they will be generated.

2.3.4 Labeling and Identification

Indicate how an individual product is identified and labeled, including PDS labels and any other labels or header information. Details about label and header formats should be specified later in section 6. Each individual product should have a unique identifier; describe how this identifier is assigned. If multiple versions of a product are possible, state how they will be distinguished.

2.4. Standards Used in Generating Data Products

2.4.1 PDS Standards

State that the data products comply with Planetary Data System standards for file formats and labels. Refer to the PDS Standards Reference.

2.4.2 Time Standards

2.4.3 Coordinate Systems

2.4.4 Data Storage Conventions

Address any data storage issues, such as byte order, machine dependence, and compression.

2.5 Data Validation

Describe validation procedures applied to data products to ensure that their contents and format are free of errors. This may be a brief overview if the details are described in another document.

3. Detailed Data Product Specifications

Describe the physical organization of the data product so that the user can access and analyze the data.

3.1 Data Product Structure and Organization

3.2 Data Format Descriptions

3.3 Label and Header Descriptions

Include PDS labels and other labels and headers that describe the data.

4. Applicable Software

Describe any applicable software used to examine, display, or analyze the data products. State who produces the software and who uses it.

4.1 Utility Programs

Include image display tools, plotting tools, etc., that allow a user to examine the data.

4.2 Applicable PDS Software Tools

List the PDS software tools that may be useful in examining the data products.

4.3 Software Distribution and Update Procedures

Describe how the software is distributed and updated, and how the user can obtain it.

Appendices

A. **Glossary**
B. **Acronyms**
C. **Definitions of Data Processing Levels**
D. **Example PDS Label**
E. **Dictionary of PDS Label Keywords**

Table 5 – Example Outline for Archive Volume SIS for Instrument-Related Data

1 INTRODUCTION
- 1.1 Content Overview
- 1.2 Scope
- 1.3 Applicable Documents

2 INTERFACE CHARACTERISTICS
- 2.1 Operations Perspective
- 2.2 Volume and Size
- 2.3 Interface Medium Characteristics
- 2.4 Backup and Duplicate Copies

3 ARCHIVE FORMAT AND CONTENT
- 3.1 Format
  - 3.1.1 Disc Format
  - 3.1.2 File Formats
    - 3.1.2.1 Text Files
    - 3.1.2.2 Document Files
    - 3.1.2.3 Catalog Files
    - 3.1.2.4 Tabular Files
    - 3.1.2.5 PDS Label Files
    - 3.1.2.6 Data Files
- 3.2 Content
  - 3.2.1 Volume Set
  - 3.2.2 Directories
    - 3.2.2.1 Root Directory
    - 3.2.2.2 Calibration Subdirectory
    - 3.2.2.3 Catalog Subdirectory
    - 3.2.2.4 Document Subdirectory
    - 3.2.2.5 Gazetteer Subdirectory
    - 3.2.2.6 Geometry Subdirectory
    - 3.2.2.7 Index Subdirectory
    - 3.2.2.8 Label Subdirectory
    - 3.2.2.9 Software Subdirectory
    - 3.2.2.10 Data Subdirectories
    - 3.2.2.11 Extras Subdirectory

APPENDICES
- A. Mars Surveyor 2001 / Odyssey
- B. Mars Surveyor 2003 / Mars Exploration Rovers
- C. Mars Surveyor 2005 / Mars Reconnaissance Orbiter

Table 6 – Example of PDS Archive Volume Structure and Contents

The following is an example of the directory structure and contents of a typical PDS archive volume or volume set. Asterisks (*) indicate directories and files that are required for PDS archives. The archive may be stored on one or more physical volumes. In the event that multiple volumes are used to store a particular data set, typically all the directories are present on each volume in the volume set.

Directory	Contents
Root	*AAREADME.TXT: Introduction and general information for the user. ERRATA.TXT: Summary of anomalies found on this volume and on previous volumes in the set, if any. (Required file if there are comments or anomalies.) *VOLDESC.CAT: Description of the volume for the PDS Catalog.
*DOCUMENT	*DOCINFO.TXT: Description of files in the document directory. *VOLINFO.TXT: Overview of the mission, instruments, and the archiving process. Archive Volume SIS Data Product SISs Science Papers
CATALOG	CATINFO.TXT: Description of files in the catalog directory. CATALOG.CAT: Descriptions of mission, spacecraft, instrument and data sets for the PDS Catalog.
*INDEX	*INDXINFO.TXT: Description of files in the index directory. *INDEX.TAB: Table listing the data products on this volume and pertinent information about each one. *INDEX.LBL: PDS label describing the format of INDEX.TAB. CUMINDEX.TAB: Table listing the data products on this volume and all previous volumes in the set. CUMINDEX.LBL: PDS label describing the format of CUMINDEX.TAB.
LABEL	LABINFO.TXT: Description of files in the label directory. Additional PDS labels and include files that are not packaged with the data products.
SOFTWARE	SOFTINFO.TXT: Description of files in the software directory Software programs, utilities, or libraries used to access and/or process the data products Source code and/or algorithms Build instructions for various platforms
CALIB	CALINFO.TXT: Description of files in the calib directory Calibration files used to create the data products or needed for processing the data products.
GEOMETRY	GEOMINFO.TXT: Description of files in the geometry directory SPICE and other files needed to describe the observation geometry.
*DATA	Data products with PDS labels. A data product's label may be attached (embedded at the beginning of the file) or detached (stored in a separate file in the same directory). Typically the DATA directory is divided into subdirectories so that no more than about 100 files are present in one directory.

12 Figures

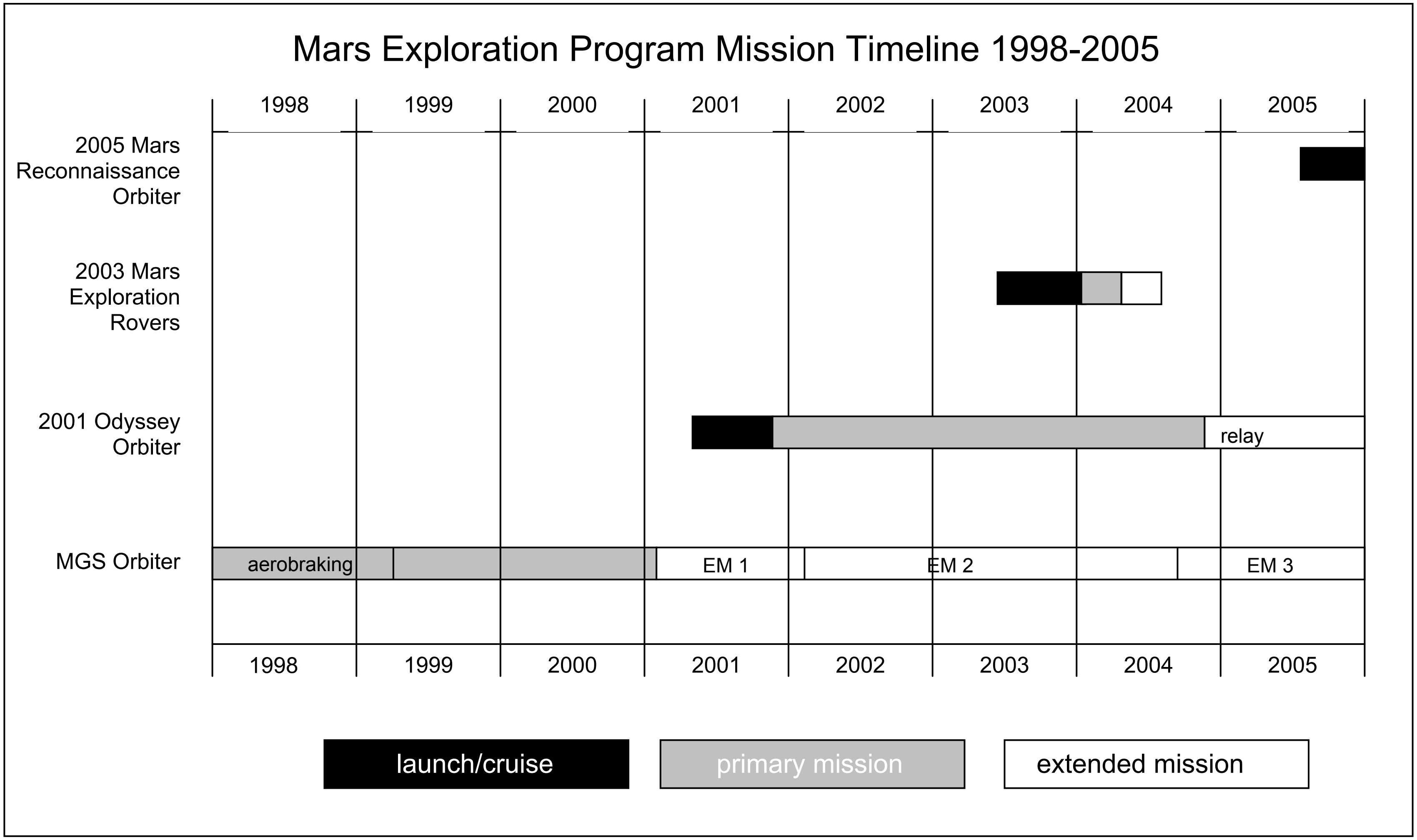

Figure 1 – Timeline for Exploration Program Projects.

Calendar years are shown.

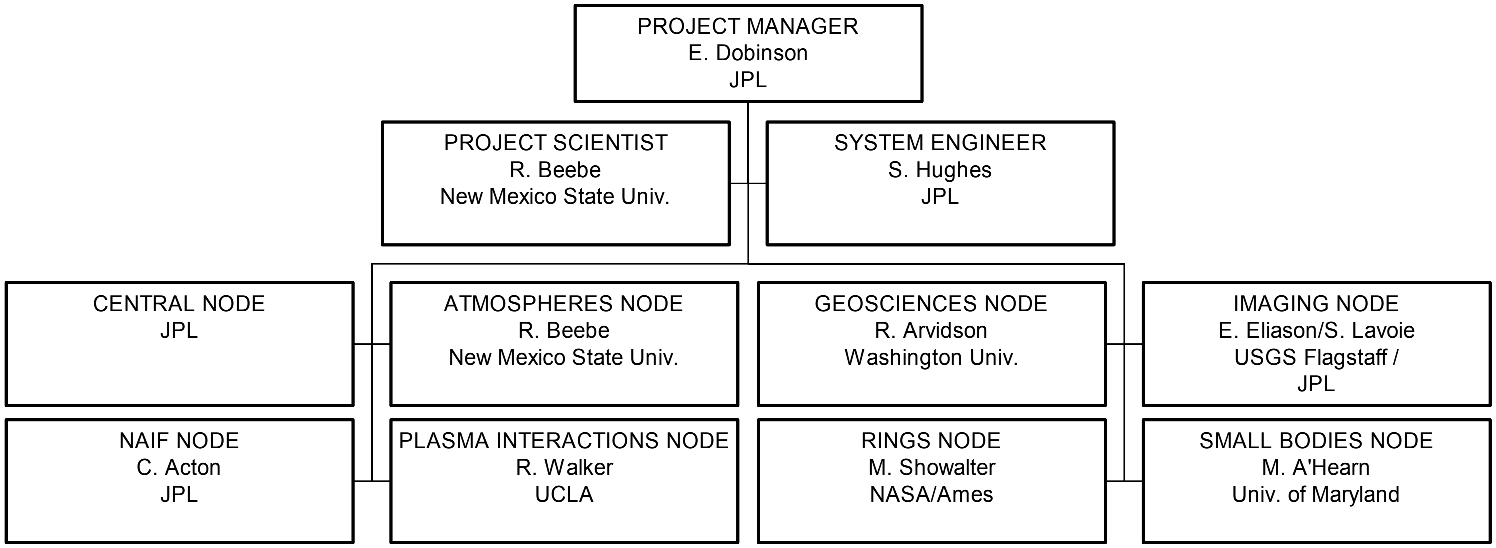

Figure 2 – Planetary Data System Organizational Structure.

Not shown are the extensive set of geographically distributed subnodes and data nodes, and the Radio Science support node at Stanford University led by R. Simpson.

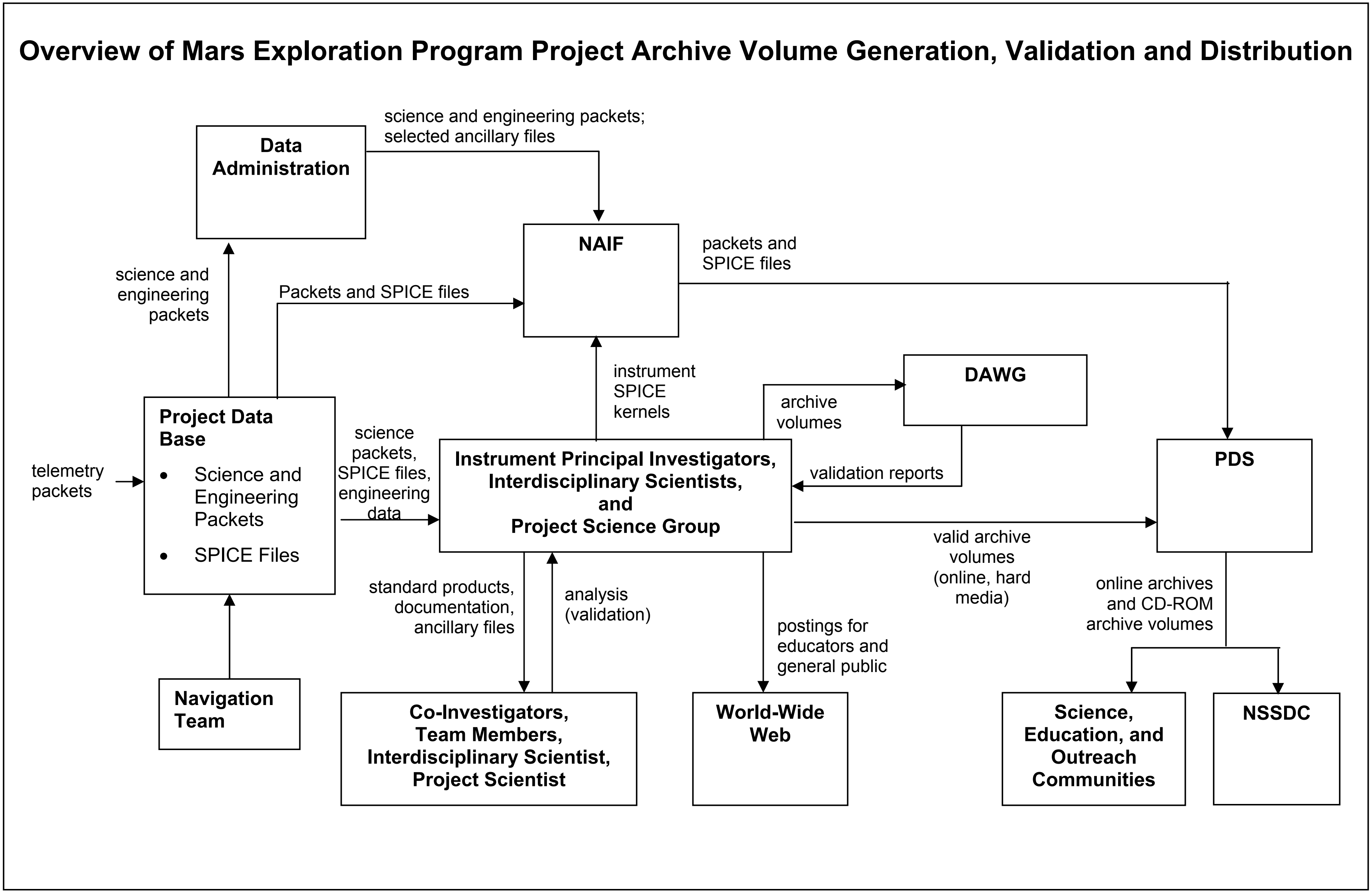

Figure 3 – Flow of Archive Volume Generation, Validation and Distribution Functions for Mars Exploration Program Projects